LIVING BEING AS INFORMED SYSTEMS: TOWARDS A PHYSICAL THEORY OF INFORMATION

Antonio León

Instituto Francisco Salinas, Salamanca, Spain

E-mail: aleonsanchez@terra.es

ABSTRACT. I propose here a new concept of information based on two relevant aspects of its expression. The first related to the undeniable fact that the expression of information modifies the physical state of its receiver. The second to the arbitrariness of such physical changes. In fact, these changes are not deducible from physical laws but from a code established arbitrarily. Thus, physical information is proposed here as the capacity of producing arbitrary changes. Once defined physical information from this physical point of view, I deduce some basic properties of informed systems. These properties (renewal, self-reproducing, evolution, diversification) are immediately recognized as the attributes most characteristic of living beings, the only natural informed systems we know. I also propose here a double evaluation of information. The former is an absolute measure of the physical effects of its expression based on Einstein's probability. The second is a functional measure based on the probability that an informed systems attain a given objective as consequence of the expression of its information.

1. INTRODUCTION

In scientific literature, especially in Biology, the word information is used with different meanings [26], although only one of them, the probabilistic (Shannon's information), has been developed in terms of what could be called a scientific theory. Among these meaning (probabilistic, algorithmic, semantic, functional, etc) almost never appears the physical aspects I will try to develop here: information as the physical capacity of the carrier systems of modifying the physical state of the receiver systems (what Kppers called the pragmatic aspect of information [37]).

As far as one can deduce from recent revisions of related subjects [20, 22, 35, 53, 65, 74, 77] there is no formal study of this undeniable physical capacity of the information carrier systems. Although some ideas more or less related to physical information can be found in the appropriate literature. Among other examples, I will point out the following:

(1) Kant's assumption that there exist in living organisms a forming force with organizational capacity, which self-propagates and cannot be explained in terms of simple mechanisms [33].

(2) Wolpert's positional information, referring to the spatial position of the system's components, which is used in determining the pattern of cellular differentiation [76].

(3) Configurational information related to the configurational microstates of the system [71].

(4) Margalef's suggestion of linking information to other physical properties of the Universe, being a property of whatever thing formed by different particles [44].

(5) Finally, Stonier work [65], where information is proposed as another essential constituent of the Universe together with matter and energy.

Stonier defines information as the capacity to organize a system or to maintain it in an organized state. But he does not take into account the arbitrariness of the established codes between the carriers and the receivers, which is proposed here as the essential feature of information. In addition, Stonier uses the term organization as equivalent to order, which is very debatable [74, 75].

The discussion that follows will begin by analyzing the inappropriateness of Shannon information and algorithmic information for the study of the physical effects of the expression of information. Immediately after it will be demonstrated that it is possible to consider certain aspects of the physical nature of the carrier systems in order to define a new concept of (physical) information based on the capacity of modifying the physical state of the receiver systems. Once defined this new concept of physical information, the physico-mathematical formalism will be applied to it, deriving as consequence the basic properties of informed systems. These properties are immediately recognized as the most characteristic attributes of living systems. In this way, physical information introduces in Physics a concept which is so strange to it as usual in biology: the concept of function [74]. Physical information reveals itself as the nexus of union between Physics and Functional Biology, which, until now, maintain a certain mutual isolation [72]. The discussion will finish by proposing a double measure, physical and functional, of information. The latter with an undoubtable evolutionary interest.

2. SHANNON INFORMATION

Shannon information theory was developed by Hartley, Shannon and Weaver in order to analyze the efficience of the transmission of electronic signals [61]. Shannon information measures the rarity of the messages formed by a certain number of symbols, each one with an occurrence probability of p_i. According to Shannon-Weaver's expression, the quantity of information per symbol (h) is given by:

$$h = -\sum_{i=1}^{k} p_i lg_2(p_i) \text{ bits} \tag{1}$$

Once the quantity h has been calculated, we can immediately calculate the amount of information (H) of any message formed by n of such symbols:

$$H = nh \tag{2}$$

Shannon information has been used with many different purposes in biology. For example, it has been applied to analyze nucleotide sequences [11, 24], biological adaptation [14], biological diversity [42, 43], ecological networks [67, 68], quantitative ecology [2], even to establish the fundamentals of biological evolution [12].

The biological success of Shannon information is probably due to the mathematical analogy between expression (1) and Boltzmann-Gibbs' expression for statistical entropy. For certain authors, this mathematical analogy makes the consideration of information as a form of entropy

(or viceversa) legitimate, and consequently, also legitimates the application of entropy laws to information, specifically the second law of thermodynamics. The validity of this supposed analogy between information and entropy, and therefore that of its corresponding biological applications, has been seriously contested [6, 18, 27, 31, 61, 66, 74, 73].

What is not debatable is the lack of semantic value of Shannon information. In effect, from equation (2) it immediately follows that all messages of equal length have the same quantity of information, i.e. the same number of bits. Shannon information has not physical significance either, since different messages with the same Shannon information will produce, in general, different effects in the receiver. Or in other words, Shannon information does not distinguishes among messages that produce different physical effects.

Let us consider a biological example. The gene that codes the cytochrome c in *Rhodopseudomonas palustris* has 342 nucleotides. Therefore there could exist 4^{432} different genes with the same Shannon information as it. Most of these genes would produce proteins with different functional abilities from that of the cytochrome c. Moreover, there could also exist a huge number of genes with different Shannon information whose products would be functionally equivalent to the considered cytochrome c. For example, the gene that codes the cytochrome c in *Pseudomonas mendocina* has 516 bits while that of *Rhodopseudomonas palustris* has 684. In consequence, from the point of view of the functional ability of proteins the number of bits of their coding genes is absolutely irrelevant.

Similar conclusions may be obtained from algorithmic information, for which the information content of a message m is equalled to the number of bits of the shortest universal computer programme that, when executed, yields as output the message m in question [79]. In this case, the lack of semantic and functional value can be immediately deduced from the enormous number of programs withe the same number of bits, but producing different outputs.

Certain authors consider that although Shannon information measurement is not related to meanings, it is sufficient to assume that meanings exist, even though they are inaccessible to us as external observers [2]. However, even when the meaning is inaccessible to us, the physical changes produced in the receiver as consequence of the expression of information can be observed and also measured. The problem is that, in spite of certain attempts [65], we do not yet have a theory of physical information. We do not even have a formal definition of physical information.

Thus, we should speculate how this situation affects our understanding of life. Since living beings are the result of the physical changes produced through the expression of the appropriate information on the appropriate receiver, can we expect to understand their physical nature without a theory of physical information? Can we expect to explain life leaving aside one of its most singular attributes. It does not seem probable. On the other hand, as we will immediately see, the consideration of the physical aspects of information does not pose serious difficulties.

3. PHYSICAL INFORMATION

As is well known, the information of a structural gene is expressed by means of the cellular apparatus of translation, provoking a physical change in its cytoplasmic receiver (basically by converting a group of free amino-acids in a peptidical chain). I will use this physical capacity

of genes as a reference in order to construct the definition of physical information. For this, it seems convenient to recall the following well established facts:

(1) The physical changes produced in the cytoplasm by the expression of genes are arbitrary, i.e. they cannot be directly deduced from physical laws, but from a code established, at least partially, in an arbitrary way.

(2) The process of translation is always the same, independently of the translated sequence.

(3) A given sequence of nucleotides always produces the same change in the cytoplasmic receiver.

(4) By simply changing the nucleotide sequence, different changes in the cytoplasm will be obtained.

(5) Genes are expressed again and again, while their nucleotide sequences do not undergo any modification due to the expression process.

It is undeniable that genes have the strange capacity of changing the physical state of their cytoplasmic receivers in many different ways, although by the same basic process. Ii is obviously about a physical capacity, since the changes produced by them are of this nature. We also know that this capacity is base on a configurational attribute of genes: their sequences of nucleotides.

From a physical point of view, attention should be paid to the arbitrariness of such changes, i.e. to the fact that the changes cannot be deduced from physical laws, they cannot be predicted without the knowledge of the genetic code. This is a truly outstanding fact that, in my opinion, states the essential difference between carrier and non-carrier information systems. In effect, whilst it is theoretically possible to express the behaviour of the latter in terms of mathematical equations deduced from the physical laws driving their evolution, it is not possible to do the same with the former: in this case we would also need to know the arbitrary code that determines the changes in the receiver.

This ability of genes, as strange as it might be, is real (genes do exist). Undoubtedly, it deserves to be somehow termed, and it seems adequate to do so by using the expression *physical information*. This is what I will propose in the following discussion, where I will try to generalize and formalize these suggestions. I will not use certain terms (like order, organization, complexity) that may appear to be appropriate in the subsequent discussion for two reasons. The first because they are not necessary to the very discussion. The second because, although they are usual terms, they are used with different meanings [74]. Consequently, I would have to justify my own selection of meanings, adding unnecessary debates. On the contrary, the terms "system" and "configuration" will be used with the maximum generality. As is usual in thermodynamics, I will use the term *system* to refer to *any arbitrarily defined part of the universe*, and the term *configuration* in order to refer to the *spatial disposition of the arbitrarily defined parts of a system*. I will also use the term multi-configurational to refer to systems with the capacity of exhibiting different configurations with the same components. We are now in the appropriate situation to propose the following:

Definition 1. *Physical information is the capacity of producing arbitrary changes.*

The systems that manifest this capacity will be termed *carriers*, while *receivers* will be those that undergo the changes. Although it it possible that both systems might be independent from one another (although with a communication channel between them), here I will deal only

with carriers and receivers integrated in only one system that I will term *informed system*. The decision is not capricious, not only because this is the case for living organisms, but also for reasons of feasibility [70] and of stability (see below the origin of information). I will use the term *expression* to refer to the set of interactions between the carrier and the receiver of which results the change of state of the latter. Finally, I will term *apparatus of translation* to the receiver components that carry out these interactions. From now on, and whenever there is no confusion, I will use the term *information* to refer to the physical information that a I have just defined.

The above definition states arbitrariness as the most outstanding characteristic of the changes produced by information. This is quite clear in human language, where the lack of physico-chemical relations between a word and its designed object is so evident [59]. The same applies to genetic language. The affinity preferences between codons and their corresponding amino acid found experimentally are not sufficient as to explain the establishment of the genetic code ([1], [15] and references therein, [35], [49] and references therein). Both the existence of homonymous codes [3, 9, 36, 58] and the distance separating the anticodon site from the site where aminoacids bind tRNA, support this experimental conclusion [35]. Although some initial prebiotic interactions, probably hydrophobic ([1] and references therein), between groups of codons and groups of aminoacids could exist, it seems clear that at least a certain arbitrariness has remained frozen in the genetic code. On the other hand, the following formal deduction of the physical properties of informed systems is independent of the grade of arbitrariness, provided that this grade may not be null.

4. Physical Properties of Informed Systems

As with all physical capacity, information must also be based on some physical attribute of its carrier system. Since arbitrary changes are determined by an arbitrary code rather than by physical laws, an attribute capable of producing arbitrary changes must comply with the following requirements:

(1) It must be capable of producing different changes in the receiver by the same type of process. Otherwise the changes would not be arbitrary but determined by physical laws.

(2) It must be multiconfigurational of wide range, otherwise it could not be the physical support of the enormous information diversity.

(3) Its different configurational states must have the ability of producing different physical changes in the receiver, while the same configurational state will always produce the same change.

(4) Arbitrariness requires that the current correspondence between the carriers' configurational states and the receiver physical changes be not the only possible.

(5) It must be a stable attribute with respect to the expression processes. If not, information would be destroyed and would have to be created continually.

These conditions imply the following

Proposition 1. *Information carrier systems must be multiconfigurational and stable in the face of the expression processes.*

The carrier configurational attribute on which information is based, would not be operative if its different configurational states do not have the opportunity of being originated approximately with the same probability. However, it is not necessary that these configurational changes have to occur in any special sequence nor in any special way. It is sufficient that they are occasionally produced. It is enough therefore, that such changes may be originated through sporadic and random interactions between the carrier and its environments. Thus, we can state the following:

Proposition 2. *Information carrier systems must have access to their different configurational states at least through random interactions with their environment.*

In natural systems, arbitrariness can only arise from the fixation of randomly selected processes. In consequence, since the very fixation is an irreversible process, it can only occur in removed from thermodynamic equilibrium systems. Therefore, informed systems must be produced and maintained away from equilibrium. This circumstance requires the continuous supply of energy, which in turn demands that informed systems have to be of a non isolated nature, with the ability of maintaining energetic interchanges with their environment. In addition to import energy the maintenance of a non equilibrium steady state also requires that the system export to its surroundings all the entropy that its irreversible processes inevitably produce. Accordingly, we can state the following:

Proposition 3. *Informed systems must be of a non isolated nature, able to maintain their non equilibrium states by interchanging energy and entropy with their environments.*

This necessity of interchanging energy and entropy involves the nature of the environment in the maintenance of informed systems. As we will see next, some of the most interesting properties of informed systems are derived from this circumstance. In order to deduce them, the following discussion about the stability of removed from equilibrium systems is necessary.

Consider any system with the ability stated in Proposition 3. The phenomenological relationships (F_i) linking the flows (J_i) of matter and energy with their generalized driving forces (X_i):

$$J_i = F_i(X_1, X_2, \ldots X_n) \tag{3}$$

are generally non lineal, and therefore very sensitive to small variations in the intensity of the forces. On the other hand, the stability conditions for non equilibrium steady states are given by:

$$\delta^2 S < 0 \tag{4}$$

$$\frac{d(\delta^2 S)}{dt} \geq 0 \tag{5}$$

where $\delta^2 S$ denotes the excess of entropy [25]. Internal fluctuations around a non equilibrium steady state arise in the system with a probability:

$$P = e^{\delta^2 S/2K} \tag{6}$$

where K is Boltzmann's constant. While conditions 4 and 5 hold, all small fluctuations in the intensity of the forces are damped down and the system recovers its steady state. Otherwise, any small fluctuation can be amplified and the system may spontaneously evolve to another macroscopic state, including the equilibrium one.

Equation 5 represents the dynamic factor's contribution to stability, such as the rate at which matter and energy are supplied to the system. Because of environmental fluctuations, variations of these energetic material supplies are unavoidable. Some of these variations may be the cause for which equation 5 no longer holds, and consequently the cause for which the system becomes unstable, sensitive to small fluctuations. Let $p_i(t)$ be the probability for an environmental fluctuation to occur during a time t causing a change in the flow J_i that the system decays to equilibrium. The fluctuating behaviour of nature allow us to state that:

$$p_i(t) > 0 \tag{7}$$

$$\frac{d(p_i(t))}{dt} > 0 \tag{8}$$

Hence, the self-maintaining probability $P(t)$ of an informed system sustained at the expense of n of these matter and energy flows may be written as:

$$P(t) = \sum_{i=1}^{n} [1 - p_i(t)] \tag{9}$$

and according to (8):

$$\frac{dP(t)}{dt} = -\sum_{i=1}^{n} \frac{dp_i(t)}{dt} \sum_{j=1}^{n} (1 - p_{j,\, j \neq i}(t)) < 0 \tag{10}$$

i.e. $P(t)$ decreases with time, and we cannot ensure the permanence of any informed system over sufficiently long periods of time. Moreover, if $dp_i(t)/dt$ are not decreasing functions -and there is no reason to assume that they are- then it can be easily shown that $d^2P(t)/dt^2 > 0$, which means that $P(t)$ decreases exponentially with time. In conclusion, we cannot ensure the existence of natural informed systems.

The only way to solve this situation is to assume that informed systems have self-replicating ability. This is not a strange capacity: living beings are self-replicating systems, and they are not the only possible ones [32]. Let us assume the existence of x self-replicating systems at a given time t, we can write the following demographic equation:

$$x_{n+1} = rpx_n \tag{11}$$

where x_{n+1} and x_n denote the number of systems of two successive generations; r the replication rate (the number of copies per system); and p the probability that the system achieve its replication. The maintenance of the number of systems of the group requires that:

$$x_{n+1} = x_n \tag{12}$$

and therefore that:

$$rp = 1 \tag{13}$$

whence:

$$r > 1 \tag{14}$$

since $p < 1$. Thus, we can state the following:

Proposition 4. *It is only possible to ensure the existence of informed systems of a self-replicating nature with a replication rate greater than 1.*

Evidently, self-replicating requires the use of environmental materials for the production of copies. That is to say, self-replicating systems must be open systems, being able to maintain exchanges not only of energy but also of matter wit their surroundings. From now on, I will deal only with these self-replicating informed systems.

Let us term generation renewal time (t_g) as the time lapsed between two successive replications of the components of a group of informed systems. Since environmental fluctuations occur with a high irregular frequency, not all the interval t_g will be equally favorable for replication. In other words, the replication probability depends on the considered period of time. Let

$$x_{n+1} = rpx_n \tag{15}$$
$$rp = 1 \tag{16}$$

be the demographic equations of a group of informed systems, p being referred to any period of time t_g. At less favorable periods, the replication probability p_L will be smaller than p and, consequently, the number of systems will decrease:

$$x_{n+1} = rp_L x_n < x_n \tag{17}$$

since

$$rp_L < rp = 1 \tag{18}$$

At more favorable periods, we will have the opposite situation ($p_M > p$) and the number of systems increases:

$$x_{n+1} = rp_M x_n > x_n \tag{19}$$

since now

$$rp_M > rp = 1 \tag{20}$$

If the recovery is not sufficient the group may be extinguished after a seres of unfavorable periods. Therefore the only groups which ensure their maintenance are those for which $rp > 1$ most of the time. But in this type of system a new class of interaction appears. In fact, let N_0 be the initial number of informed systems of one of these groups. At the end of k favorable consecutive generations the number of elements of the group will be:

$$N_k = (rp)^k N_0; \ rp > 1 \tag{21}$$

The above exponential growth in the number of systems would also demand an exponential growth of the environmental material and energetic resources, which is impossible. Thus, the limitation of environmental resources imposes restrictions in the number of systems of the group, and therefore, the appearance of demographic competitive interactions: in many generations more copies are formed than can be sustained by the environment. Consequently, we come to:

Proposition 5. *Informed systems are liable to undergo extinctions and demographic competitive interactions.*

Let us consider now any group of informed systems. In general, the carriers of different systems of the group will exhibit different configurational states that, we can assume, have been produced by simple random interactions with their environment. Naturally, the only possible configurational states of the carriers will be those which produce changes in the receivers compatible with their necessity for maintenance and replication. But not all of them will be equally efficient in relation to such activities. Let us assume, for the sake of simplicity, that our group is formed by two types of systems, A and B, so that their respective replication probabilities verify:

$$p_B < p_A \tag{22}$$

Let A_k and A_{k+1} be the number of A-systems in two successive generations. Its relative variation rate (d_A) can be written as:

$$d_A = \frac{A_{k+1} - A_k}{A_k}; \ A_k > 0 \tag{23}$$

and taking into account that:

$$A_{k+1} = rp_A A_k \tag{24}$$

we can write:

$$d_A = \frac{rp_A A_k - A_k}{A_k} \tag{25}$$

$$= rp_A - 1 \tag{26}$$

Similarly, for B-systems we come to:

$$d_B = rp_B - 1 \tag{27}$$

Finally, from (22) it follows that:

$$d_B = rp_B - 1 < rp_A - 1 = d_A \tag{28}$$

which means the relative increase of A-systems with respect to B-systems. And consequently, the increase of the average replication probability of the whole group. Note that this increase does not depend on the demographic state of the group, which may be stationary or variable. Therefore, we can state the following:

Proposition 6. *Informed systems continually increase their replication probability, until reaching the maximum value compatible with their environmental conditions.*

The above proposition states the existence of a directional, evolutionary process in informed systems. In reality, this proposition is a physical and formalized version of Darwin's Principle of Natural Selection.

Finally, we will consider the anisotropy of the natural environment. That is to say, the qualitative and quantitative differences in its properties. These differences allow us to consider the natural environment as formed by a set of mosaics differentiated by some properties which are significant from the point of view of the maintenance and replication of informed systems. Now, all we have to do is to apply Proposition 6 to the systems of each one of these mosaics, thus deducing that he changes which contribute to increase the replication probability of the systems under the particular conditions of each mosaic are to be propagated. We thus have the following:

Proposition 7. *The environmental anisotropy produces the diversification of informed systems.*

It seems appropriate now to make a brief résumé and an evaluation of the results obtained until this point. We have proposed a definition of physical information and we have applied to it the physco-mathematical formalism. By doing so, we have deduced the most basic physical properties of informed systems, that can be resumed in the following way. Informed systems:

(1) have multiconfigurational carriers, which undergo changes in their configurational states.
(2) are open systems, capable of interchanging matter, energy and entropy with their environment.
(3) are self-replicating systems, with a replication rate greater than one.
(4) are exposed to undergo extinctions and competitive interactions.
(5) suffers evolutionary processes.
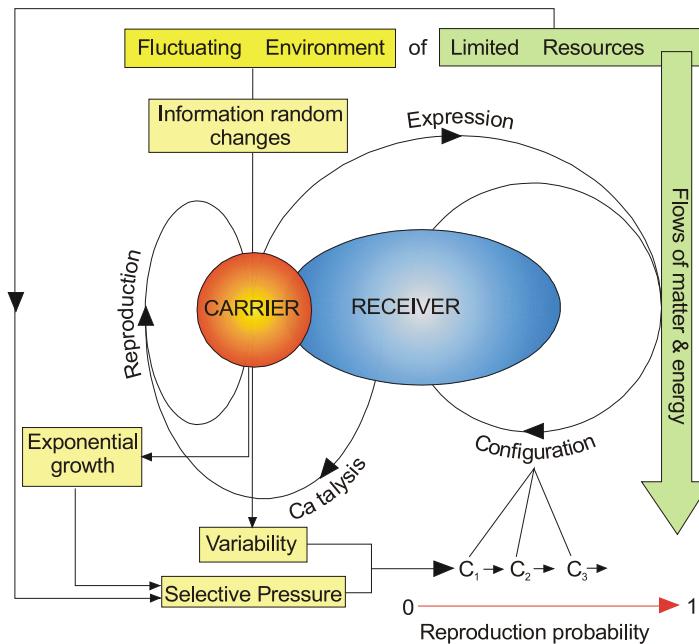(6) suffers diversification processes.

FIGURE 1. Living beings as informed system: operating and evolution

It is obvious to recognize in these properties the most significant attributes of living beings, which are the only natural systems we know of (Figure 1). Although no new property of living beings has been discovered here, the method used to obtain them is a significant novelty: the physico-mathematical formalism applied to a definition of information that consider it as a physical entity rather than a theoretical one.

In spite of the conflict that may result from proposing the existence of a new physical entity, in the case of information the only thing is to do it. Living organisms exist and, in effect, their genes show the capacity that we have defined as physical information. In addition, the proposed definition is very close to that of energy (the capacity of producing changes), the only significant difference being the arbitrariness of the changes produced by information in the face of the physical determinism of the energetic changes. Information could be really considered as a sort of energy, which arises in system removed from thermodynamic equilibrium, as a consequence of the fixation of certain randomly selected processes. To compensate, this new physical concept will allow us that living organisms no longer be these historical objects so strange to physics [10, 51, 55, 57, 56]. It will also allow the development of the physico-mathematical formalism necessary in my opinion, to end once and for ll the controversy permanently rooted in the very foundations of evolutionary biology [19, 28, 40, 63, 69]. In addition to the qualitative

explanations, we will immediately see that it is also possible to obtain quantitative results related to the physical and functional effects of information.

5. Information and Functional Activity

The concept of function is related to the activity of certain systems such artificial machines or living beings. It can be considered as the objective or purpose that the system has to attain. Although it can also be considered as the activity developed in order to attain the objective in question (functional activity). Here, I will use the term with the maximum generality: a *function* will be *any arbitrarily defined purpose*. In natural nonliving objects, there is no sense is speculating about their objective or purposes, since they are nothing more than the products of the evolution of the universe under the driving force of physical laws. For this reason, the concept of function does not belong to the lexicon of physics [74]. As we will see, in the case of an informed system there are objectives or functions for which it makes sense to speculate.

In effect, we already know that informed systems have arbitrary characteristics (which are produced by the joint action of the random modifications of the carriers and the arbitrary inter-actions between the carriers and the receivers). Although these characteristics are not deducible from physical laws, they are the same laws which determine their permanence or extinction. Because they are those which govern the behavior of the natural environment where informed systems try self-maintaining and self-replicating, and those which impose restriction to the very physico-chemical processes involved in such activities of maintenance and replication. One can say that physical laws put to test the arbitrary attributes of informed systems. It makes sense, therefore, to ask whether a given informed system will reach the objective of its own replica-tion. The answer is not obvious, and the lack of obviousness justifies the physical consideration of the concept of function and that of functional activity. So, informed systems introduce in Physics these concepts so typically biological.

Naturally, each function will be attained from a certain state of the system, which in turn will be the result of the expression of its information. In this way it will be possible to evaluate infor-mation from the point of view of the functional ability of the system. From a practical point of view, it may be interesting to decompose a complex function (as presumably is self-replicating) in an ordered series of subfunctions, each one of them attained from a certain state of the sys-tem, which in turn results from the expression of a certain "fragment" of information. For this to be possible, the system has to be able -as occurs in living beings- to express an ordered series of fragments of information, which does not add any new difficulty of its functioning. It is sufficient to consider -again as in living organisms- the appropriate interactions between certain components of the receiver and certain special fragments of information (metainfor-mation) [7, 41, 54], which, finally, will determine the expression of the adequate fragment of information.

6. Physical and Functional Measurements of Information

As I have repeatedly said, the expression of information modifies the state of its receivers sys-tems. Next, I will propose a way of measuring the intensity of these modifications. Statistical thermodynamics provides us with the appropriate grounds for such an objective. In effect, as

is well known, a change in the state of a system involves a change in the number of its compatible microstates. Thus we can estimate the physical effects of information in terms of the variation in the number of such microstates suffered by the system because of the expression of its information.

Consider two macrostates of an informed system. The first that of thermodynamic equilibrium, the second the resulting from the expression of its information. Let W_e and W_i be the number of microstates which are respectively compatible with such macrostates. I will define the measurement of physical information, Ω, as:

$$\Omega = \ln\left(\frac{W_e}{W_i}\right) \tag{29}$$

The quantity Ω is a dimensionless measurement that represents the decreasing rate in the number of microstates caused by the expression of information. In other words, it represents the degree of removal from thermodynamic equilibrium caused by information (it could be said to be the degree of "informization"). It is immediate to express Ω in terms of entropy variations. Indeed, according to Boltzmann's equation we can write:

$$S_e = K \ln W_e \tag{30}$$

$$S_i = K \ln W_i \tag{31}$$

where S_e and S_i denote the system entropy at the considered conditions, and K is Boltzmann's constant. From (30) and (31) we immediately come to:

$$\Delta S = S_i - S_e = K \ln\left(\frac{W_i}{W_e}\right) \tag{32}$$

hence:

$$\frac{W_e}{W_i} = e^{(-\Delta S/K)} = P^{-1} \tag{33}$$

where P is Einstein's probability, which is the probability for a fluctuation to arise in a system at equilibrium, originating the change ΔS in the entropy of the system. According to (33) we can write (29) as:

$$\Omega = \frac{-\Delta S}{K} \tag{34}$$

or:

$$\Omega = -\ln P \tag{35}$$

since Ω is a dimensionless quantity, we can define ay arbitrary unit of measurement. For instance, we can define the physical information unit (pit) as the quantity of information for which $W_i = 1/2W_e$. That is:

$$1 \text{ pit } = \ln\left(\frac{W_e}{(1/2)W_e}\right) = \ln 2 \tag{36}$$

Therefore (34) and (35) can be rewritten as:

$$\Omega = -\frac{\Delta S}{K}\ln 2 \text{ pits} \tag{37}$$

and:

$$\Omega = -\lg_2 P \text{ pits} \tag{38}$$

Note that the physical meaning of Ω is not debatable: it has been defined explicitly in terms of entropy variations.

As is to be expected, the values of P in living organisms are extremely small. For instance, in the case of *Escherichia coli* the probability of a living microstate is in the order of $10^{-10^{11}} = 10^{-100000000000}$ [50]. The changes in entropy ΔS can be estimated, at least for single unicellular organisms, by micro-calorimetry [39] (an estimate of $-3.35 \times 10^{-12} JK^{-1}$ in the variation of entropy during the process of the formation of a cell was found empirically by Lurié and Wagensberg; accordingly, the quantity Ω involved in the process would be of 3.5×10^{11} pits).

In addition to measuring the global effects of information, we can measure the partial effects of the expression of a given fragment of information. We can use the same definition (29):

$$\Omega_{12} = \ln\left(\frac{W_1}{W_2}\right) \tag{39}$$

W_1 and W_2 now denoting the number of microstates of the receiver before and after, respectively, the expression of the considered information.

The measure of physical information we have just defined, although interesting from the theoretical point of view, has no functional value. In fact, due to the arbitrariness of the changes that information produces in its receiver, many of them can produce the same variations in the number of microstates (the same Ω) but very different functional abilities in the receiver. On the other hand, it is easy to establish a functional measurement of information, which could also be interesting form the semantic point of view. For this we need to know the function for which information is evaluated. In fact, let u be any function defined for an informed system, and let p be the probability of achieving this function once expressed the information I on a certain receiver state. I will define the functional measure, ϕ of I as:

$$\phi = f(p) \tag{40}$$

where f is a type C^2 (continuous function whose first and second derivatives are also continuous) real function defined on $[0, 1]$ such as:

$$f(p) = 0 \Leftrightarrow p = 0 \tag{41}$$

$$f'(p) > 0, \ \forall p \in [0, 1] \tag{42}$$

which states that ϕ increases strictly as the probability p of attaining the function in question increases. Due to the extremely complex non-lineal dynamics of nature, it seems reasonable to assign an infinite value to the measurement ϕ which corresponds to the information capable of assuring the attainment of the function u in any possible environmental conditions ($p = 1$). Thus, we might also require:

$$f(1) = \infty \tag{43}$$

$$f''(p) > 0 \tag{44}$$

Any mathematical function satisfying (41) and (42) could be used to measure functional information. I propose the following:

$$\phi = \frac{p}{1 - p} \tag{45}$$

which also satisfies (43) and (44). The quantity ϕ is a dimensionless measurement too, thus we can define any arbitrary unit for it. For instance we can define the functional information unit (fit) as the quantity ϕ for which $p = 1/2$. That is to say:

$$1 \text{ fit } = f(1/2) = \frac{1/2}{1 - 1/2} = 1 \tag{46}$$

therefore:

$$\phi = \frac{p}{1 - p} \text{ fits} \tag{47}$$

From the empirical point of view, the functional measurement of information results are much more interesting than the physical one. For living beings, beside the basic function of self-reproducing, many other biological functions can be used to evaluate their genetic information in the quantitative, probabilistic terms we have just defined.

7. THE ORIGIN OF INFORMATION

Although it is not an objective of this paper, it seems appropriate to make some considerations on the origin of information. As is to be expected, this question is equivalent to that of the origin of life from the perspective of physical information. I will simply show that this new perspective does not add any new problem.

Let us begin by asking ourselves about the possible nature of the carrier and the receiver systems. The same quantum-mechanical reasons from which Schrödinger derives the molecular nature of its "aperiodic crystals" [60], will serve us to deduce the molecular nature of the carrier systems. In effect, only molecules could reach the size, variety and stability required by a multiconfigurational system. Carrier molecules have to be able to acquire different equiprobable configurations, and hence they have to be formed by at lest two different components with the appropriate binding ability so as to constitute the carrier in question. These components cannot be individual atoms simply because different atoms will bind to one another with different bonds, which is incompatible with the required equiprobability of the resulting arrangements. The components must be molecules able to form at least two equivalent chemical bonds allowing equiprobable configurations. That is to say, they have to be copolymers made up of two or more different types of monomers, the sequence of such monomers being the multiconfigurational attribute upon which the carrier capacity we have termed physical information is based.

Consequently, the receivers must also consist of molecules, which should be sensitive to the different carrier sequences, in the sense that their different states must be arbitrarily determined by the sequence of carrier monomers, rather than by any type of spatial complementarity or physico-chemical affinity between them. One way for this to be possible is that the receiver be also formed by copolymers. In these circumstances the process of expression would consist of translating the sequence of carrier monomers into a sequence of receivers monomers through an arbitrary code and an appropriate molecular apparatus of translation. Thus, the first problem we have to deal with is about the origin of such monomers and polymers. This is a classical problem in protobiogenesis on which important researches exist, whose conclusions can be reviewed in any of the numerous text devoted to the subject (for instance [13], [17], [22], [47], [23], [30], [52], [53]). In summarizing, we can recall that such molecules could have been formed, with more or less probability, on the early Earth.

Once the polymers were available, the appropriate interactions among them had to settle down so as to become living (informed) systems. We can separate these interactions into two different groups: catalytic interactions driving the process of carrier replication; and informed interactions consisting in the translation of the carrier monomer sequence into the receiver one. Both types of interactions are being studied intensely at present. There are several models for the first type of interaction, among them the model of hypercycles [21, 20] and the model of autocatalytic sets of catalytic polymers [35]. The main problem of the second type of interaction is the origin of the genetic code, for which there are some interesting suggestions, such as those of Crick [16] and Bedian [8]. Many of these investigations reveal the astonishing emergent properties of systems subjected to iterative processes, including the emergence of stabilizing attractors in systems whose components suffer iterative cycles of mutual interactions [4, 5, 29, 34, 38, 45, 46, 62, 64, 78]. All these suggest an origin of living (informed) system more expected [35] than miraculous [48]. Thus we can conclude that the origin of life, and consequently that of physical information, although still unsolved, is a scientific question rather than a metaphysical one.

8. Conclusions

The consideration of information as the capacity of producing arbitrary changes is not only necessary (genes do exist and they show just that ability) but also productive: it allows to deduce in a formal way the most significant attributes of living beings. Thus, information introduces in Physics a property so strange to the non-living world as genuine in living ones: arbitrariness. By the simple consideration of this property, life can be explained in physical terms. As we have just seen, present day researches on the origin and evolution of life suggest more and more that living (informed) systems, rather than a miracle of chance, are expected consequences of physical laws. Life seems to be nothing more than random and arbitrariness under the effects of physical laws.

We have also seen how the physical effects of information can be doubly evaluated. On the one hand it is possible, at least theoretically, to estimate the intensity of the changes suffered by the receiver in terms of entropy variations. On the other hand, it is also possible to evaluate in probabilistic terms its functional capacity. That is to say, its ability to reach a given objective. The proposed functional measurement of information is of practical use and may be of evolutionary interest, since it allows us to evaluate quantitatively genetic information.

Acknowledgement

References

[1] J. A. Aguilera, *Luces y sombras sobre el origen de la vida*, Mundo Científico (1993), no. 136, 508 – 519.

[2] H. Atlan, *Information theory and self organization in ecosystems*, Canadian Bull. of Fisheries and Aquatic Sciences **183** (1985), 187 – 199.

[3] F. J. Ayala and J. A. Kiger Jr., *Genética moderna*, Omega, Barcelona, 1984.

[4] P. Bak and K. Chen, *Criticalidad auto-organizada*, Investigación y Ciencia (Scientifc American) (1991), no. 174, 18 – 25.

[5] P. Bak, K. Chen, and M. Creutz, *Criticality In the game of life*, Nature (1989), no. 342, 780 – 782.

[6] J. M. Barry, *Informational Dna: a useful concept?*, Trends Biochem. Sci. (1986), no. 11, 317 – 318.

[7] M. Beato, *Interacciones entre proteínas reguladoras y Adn*, Investigación y Ciencia (Scientifc American) (1991), no. 174, 6 – 18.

[8] V. Bedian, *The possible role of assignment catalyst In the origin of the genetic code*, Origin of Life (1982), no. 12, 181 – 204.

[9] M. J. Berry, L. Banu, Y. Chen, J. Mandel, J. D. Kieffer, J. W. Harney, and P. R. Larsen, *Recognition of Uga as a selenocysteine codon In Type I deiodinase requires sequences In the 3' untranslated region*, Nature (1991), no. 353, 273 – 276.

[10] D. Bohm, *Algunas observaciones sobre la noción de orden*, Hacia una Biología Teórica (C. H. Waddington, ed.), Alianza, Madrid, 1976, pp. 215 – 277.

[11] D. R. Brooks, D. D.Cumming, and P. H. Leblond, *Dollo's Law and the second law of thermodynamics*, Entropy, Information and Evolution (B. H. Weber, D. J. Depew, and J. D. Smith, eds.), MIT Press, Cambridge, 1988, pp. 189 – 224.

[12] D. R. Brooks and E. O. Wiley, *Evolution as entropy*, University of Chicago Press, Chicago, 1988.

[13] C. Chyba and C.Sagan, *Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origin of Life*, Natrue (1992), no. 355, 125 – 132.

[14] M. Conrad, *Adaptability*, Plenum Press, New York, 1983.

[15] F. H. C. Crick, *The Origin of the Genetic Code*, J. Mol. Biol. (1968), no. 38, 367 – 379.

[16] F. H. C. Crick, S. Brenner, A. Klug, and G. Pieczenik, *A speculation on the origin of protein synthesis*, Origin of Life (1976), no. 7, 368 – 397.

[17] W. Day, *Génesis en el planeta Tierra*, H. Blume, Madrid, 1986, (Genesis on the Planet Earth, Yale University Press, 1986).

[18] D. J. Depew and B. H. Weber, *Consequences of non-equilibrium Thermodynamics for Darwinism*, Information and Evolution (D. J. Depew, B. H. Weber, and J. D. Smith, eds.), MIT Press, Cambridge, 1988, pp. 317 – 354.

[19] G. A. Dover, *Universal Darwinism*, Nature (1992), no. 360, 505.

[20] M. Eigen, *Steps towards life*, Oxford University Press, New York, 1992.

[21] M. Eigen and P. Schuster, *The Hypercycle: A Principle of Natural Self-organization*, Springer-Verlag, New York, 1979.

[22] R. F. Fox, *Energy and the Evolution of Life*, W. H. Freeman, New York, 1988.

[23] S. W. Fox and K. Dose, *Molecular Evolution and the Origin of Life*, W. H. Freeman, San Francisco, 1972.

[24] L. Gatlin, *Information Theory and the Living System*, Columbia University Press, New York, 1972.

[25] P. Glansdorff and I. Prigogine, *Structure, Stabilité et Fluctuations*, Mason, Paris, 1971.

[26] H. Hacken, *Misinformation*, Nature (1993), no. 362, 509.

[27] A. Hariri, B. Weber, and J. Olmsted III, *On the validity of Shannon information calculations for molecular biological sequences*, J. Theor. Biol. **147** (1990), 235 – 254.

[28] C. Hellier, *Darwinism not tautological*, Nature **351** (1991), 600.

[29] Douglas R. Hofstadter, *Los atractores extraños son configuraciones matemáticas en equilibrio entre orden y caos*, Investigación y Ciencia (Scientifc American) (1982), no. 64, 103 – 113.

[30] J. Horgan, *En el principio ...*, Investigación y Ciencia (Scientifc American) (1991), no. 175, 80–90.

[31] J. Olmsted III, *Observation on evolution*, Entropy, Information and Evolution (B. H. Weber, D. J. Depew, and J. D. Smith, eds.), MIT Press, Cambridge, 1988, pp. 243 – 262.

[32] J. Rebeck Jr., *Moléculas sintéticas autoreplicantes*, Investigación y Ciencia (Scientifc American) (1994), 68 – 75.

[33] Immanuel Kant, *Crítica del juicio*, Espasa Calpe, Madrid, 1984.

[34] S. A. Kauffman, *Anticaos y adaptación*, Investigación y Ciencia (Scientifc American) (1992), no. 184, 46 – 53.

[35] _____ , *The Origins of Order*, Oxford University Press, New York, 1993.

[36] Y. Kawaguchi, H. Honda, J. Taniguchi-Morimura, and S. Iwasaki, *The codon Cug is read as serine in a asporigenic yeast Candida cylindracea*, Nature **341** (1989), 164 – 166.

[37] B. Küppers, *Information and the origin of life*, MIT Press, 1990.

[38] Antonio León, *Coevolution: New Thermodynamic Theorems*, Journal of Theoretical Biology **147** (1990), 205 – 212.

[39] D. Lurié and J. Wagensberg, *Termodinámica de la evolución biológica*, Investigación y Ciencia (Scientifc American) (1979), no. 30, 102 – 113.

[40] J. Maddox, *Is Darwinism a thermodynamic necessity?*, Nature **350** (1991), 653.

[41] J. R. Manak and P. S. Mattew, *Able assistants for homeodomain proteins*, Current Biology **3** (1993), 318 – 320.

[42] R. Margalef, *Perspectivas de la Teoría Ecológica*, Blume, Barcelona, 1981.

[43] R. Margalef and E. Gutierrez, *How to introduce connectance in the frame of an expression for diversity*, Am. Nat. **121** (1983), 601 – 607.

[44] Ramón Margalef, *La Biosfera, entre la termodinámica y el juego*, Omega, Barcelona, 1980.

[45] Robert M. May, *Simple mathematical models with very complicated dynamics*, Nature **261** (1976), 459 – 467.

[46] _____ , *Le chaos en biologie*, La Recherche (1991), no. 232, 588 – 598.

[47] V. García Molina and J. A. Aguilera, *Y la Tierra palpitó*, H. Blume, Madrid, 1985.

[48] J. Monod, *Le Hazard et la Nécessité*, Ed. Du Deuil, Paris, 1970.

[49] D. Moras, *Synthetases gain recognition*, Nature **344** (1990), 195 – 197.

[50] J. H. Morowitz, *Energy Flow in Biology*, Academic Press, London, 1968.

[51] G. Nicolis and I. Prigogine, *Exploring Complexity*, W. H. Freeman, New York, 1989.

[52] L. E. Orgel, *Los orígenes de la vida*, Alianza, Madrid, 1975.

[53] _____ , *Origen de la vida sobre la Tierra*, Investigación y Ciencia (Scientifc American) (1994), no. 219, 47 – 53.

[54] S. E. V. Phillips, *Saddling up the Tata box*, Current Biology **3** (1993), 112 – 114.

[55] I. Prigogine, *¿tan solo una ilusión?*, Tusquets, Barcelona, 1983.

[56] I. Prigogine and I. Stengers, *Entre el Tiempo y la Eternidad*, Alianza, Madrid, 1990.

[57] Ilya Prigogine and Isabel Stengers, *Entre el tiempo y la eternidad*, Alianza, Madrid, 1990.

[58] J. Rennie, *Los nuevos giros del Adn*, Investigación y Ciencia (Scientifc American) (1993), no. 200, 68 – 75.

[59] F. Saussure, *Curso de Lingüística General*, Losada, Buenos Aires, 1973.

[60] E. Schrödinger, *What is Life*, Cambridge University Press, 1967.

[61] C. E. Shannon and W. Weaver, *The mathematical theory of Communication*, University of Illinois Press, Urbana, 1949.

[62] K. Sigmund, *Game of Life: Explorations in Ecology, Evolution and Behaviour*, Oxford University Press, New York, 1993.

[63] J. Maynard Smith, D. Bohm, M. Green, and C. H. Waddington, *El status del neodarwinismo*, Hacia una Biología Teórica (C. H. Waddington, ed.), Alianza, Madrid, 1976, pp. 295 – 324.

[64] I. Stewart, *Does God Play Dice? The New Mathematics of Chaos*, Penguin Books Ltd, Harmondsworth, 1989.

[65] T. Stonier, *Information and the Internal Structure of the Universe*, Springer-Verlag, London, 1990.

[66] C. I. J.M. Stuart, *Bioinformational equivalence*, J. Theor. Biol. **113** (1985), 611 – 636.

[67] R. Ulanowicz, *Growth and Development*, Springer-Verlag, New York, 1986.

[68] J. Wagensberg, J. García, and R. V. Solé, *Connectivity and information transfer in flow networks: two magic numbers in ecology?*, Bull. Math. Biol. **52** (1990), 733 – 740.

[69] M. E. Weale, *Darwinism not tautological*, Nature **351** (1991), 600.

[70] B. H. Weber, D. J. Depew, C. Dyke, S. N. Salthe, E. O. Schneider, R. E. Ulanowicz, and S. J. Wcken, *Evolution in a thermodynamic perspective: an ecological approach*, Biology and Philosophy **4** (1989), 373 – 405.

[71] J. S. Wicken, *A thermodynamic theory of evolution*, J. Theor. Biol. **87** (1980), 9 –23.

[72] _____ , *Thermodynamics and the conceptual structure of evolution theory*, J. Theor. Biol. **117** (1985), 363 – 383.

[73] _____ , *Entropy and Information: suggestion for common language*, Philosophy of Science **54** (1987), 176 – 193.

[74] _____ , *Evolution, Thermodynamics and Information*, Oxford University Press, New York, 1987.

[75] _____ , *Can the information content of biological systems be quantified?*, Syst. Res. **6** (1989), 133 – 142.

[76] L. Wolpert, *La formación de modelos en el desarrollo biológico*, Investigación y Ciencia (Scientifc American) (1978), no. 27, 78 – 79.

[77] H. P. Yockey, *Information Theory and Molecular Biology*, Cambridge University Press, Cambridge, 1992.

[78] M. Zeleny, *Self-organization of Living Systems*, Int. J. Gral. Syst. **4** (1977), 13 – 28.

[79] W. H. Zurek, *Thermodynamic cost of computation, algorithmic complexity and the information metric*, Nature **341** (1989), 119 – 124.